

# 동영상에서의 세 대칭적 움직임 벡터를 이용한 카메라 회전각 분석

## Analysis of Camera Rotation Using Three Symmetric Motion Vectors in Video Sequence

문성현, 박영민, 윤영우 .

Sung-Hun Moon, Young-Min Park and Young-Woo Yoon

### 요약

본 논문은 단안 영상열에서 기하학적으로 대칭인 세 점의 움직임 벡터의 특수한 관계를 이용하는 카메라 움직임 추정기술을 제안한다. 제안하는 기술은 특징점과 외극기하적 제한조건을 사용하지 않고, 카메라 회전에 의해 야기되는 움직임 벡터와 그 특성을 이용한다. 외극기하적 제한조건을 사용하는 경우에는 E-행렬이나 F-행렬의 계산을 위해 계산시간이 많이 소요되는 수치적 해법이나 반복법을 사용해야 하지만, 제안한 기술에서는 선형방정식을 움직임 벡터에 적용하여 상하회전각, 좌우회전각, 정면회전각과 배율을 한번에 추정한다. 그리고 배경판별식을 고안하여 배경영역만을 계산에 반영하기 때문에 계산이 보다 정확하고 MPEG-4 요구조건을 수용하기에 충분히 빠르다. 다양한 영상에 대한 실험결과를 통하여 제안된 기술의 타당성과 광범위한 응용가능성을 입증한다.

### Abstract

This paper proposes a camera motion estimation technique using special relations of motion vectors of geometrically symmetrical triple points of two consecutive views of single camera. The proposed technique uses camera-induced motion vectors and their relations other than feature points and epipolar constraints. As contrast to the time consuming iterations or numerical methods in the calculation of E-matrix or F-matrix induced by epipolar constraints, the proposed technique calculates camera motion parameters such as panning, tilting, rolling, and zooming at once by applying the proposed linear equation sets to the motion vectors. And by devised background discriminants, it effectively reflects only the background region into the calculation of motion parameters, thus making the calculation more accurate and fast enough to accommodate MPEG-4 requirements. Experimental results on various types of sequences show the validity and the broad applicability of the proposed technique.

**Key words:** camera motion estimation, global motion compensation, motion vectors

## I. 서론

카메라 움직임 추정은 물체의 깊이분석, 비디오 콘텐츠 분석, 그리고 영상분할과 압축 등 다양한 분야에 이용된다.

컴퓨터 비전에서는 2D 영상에서 3D 정보 복원을 위한 깊이분석에 카메라 움직임 매개변수가 필요하다. 멀티미디어 데이터베이스 응용 및 인식분야에서는 비디오 콘텐츠의 색인과 검색에 사용할 선명한 정지영상을 얻기 위해서 카메라 움직임 추정은 필수 불가결하다.

MPEG-2에서는 카메라 움직임 추정값을 사용하여 전역움직임을 보상하여 실제로 움직이지 않은 배경에 해당하는 매크로 블럭(macro block)의 벡터를 제거하게 되므로 압축효율을 높일 수 있다. 한편, 물체(Object)의 움직임은 대개 주인공의 행동이나 주요 사건을 나타내게 된다.

MPEG-4에서 이러한 동체를 포함한 전경을 분리할 경우에도, 카메라 움직임을 보상하면 동체의 벡터가 카메라 움직임에 의한 벡터와 실제 동체의 움직임 벡터의 합성이 아닌 순수한 동체의 움직임 벡터를 구할 수 있게 되므로 벡터를 이용한 영상분할이 보다 더 정확해진다.[1,2] 이와 같이 카메라 움직임 추정과 보상은 불필요한 움직임 벡터를 제거하거나 줄이는 데에 필수적인 전처리이며 대부분의 응용분야에서 실시간 처리성능이 요구된다.

그러나, 최근까지 연구된 대부분의 기술들에서, 방대한 계산량에 의한 고속구현의 부적합, 전경 정보의 포함에 의한 움직임 추정 결과의 부정확, 혹은 응용대상의 제한 등의 한 두 가지 문제가 발견된다.

본 논문은 영상 프레임에서 기하학적으로 대칭인 세 점의 움직임 벡터의 특수한 관계를 이용하여 카메라의 움직

임을 추정하는 기술을 제안한다. 제안된 기술은 1원1차 선형 방정식을 움직임 벡터에 단 한번 적용하여, 상하회전(tilting), 좌우회전(panning), 정면회전(rolling)과 배율변경(zooming)과 같은 여러 가지 카메라 움직임 매개변수들을 한꺼번에 산출하며, 기존 연구에서와 같은 수치해석적 계산이나 반복계산(iteration)을 사용하지 않는다. 또한 영상에서 동체에 해당하는 화소(pixel)와 배경에 해당하는 화소를 구분해내는 판별식을 사용하여, 전경을 배제하고 배경 정보만을 사용하여 카메라 움직임 추정을 한다. 따라서 계산이 보다 정확하고 MPEG-4 요구조건을 수용하기에 충분히 빠르다. 다양한 종류의 영상에 대한 실험결과를 제안된 기술의 타당성과 광범위한 응용가능성을 입증한다.

## II. 기존 연구에 대한 고찰

카메라 움직임은 상하회전, 좌우회전, 정면회전과 배율변경, 좌우이동(tracking), 상하이동(booming), 전후이동(dolly) 등으로 구성된다.

카메라 움직임을 분류해주는 구체적 특징은 움직임 필드(motion field)인데, 각각의 특정한 카메라 조작은 움직임 필드의 특정한 패턴으로 귀착된다. 이러한 사실을 기반으로 하여 Zhang[3] 등은 일련의 점진적으로 변화하는 영상열에서 좌우회전과 배율변경을 분류하는 간단하지만 효과적인 카메라 움직임 분석 기술을 개발하였다.

카메라 움직임에 대한 보다 정교하고 정량적인 시도는 공간의 점과 그에 해당하는 영상 좌표간의 변환 관계(transformation relation)를 사용한다. 여기서 변환은 좌우회전이나 배율변경이 있을 경우의 화소의 좌표 변화를 유도하는데 사용된다. 간단한 배율변경은 카메라 몸체의 이동이 없이 카메라 초점거리의 변경에 의해 야기된다. 반면에 좌우회전은 영상평면에 평행한 축에 대하여 카메라를 회전함에 의해 야기된다. 두 프레임으로부터 계산되는 움직임 필드와 반복적 알고리즘을 사용하여 카메라의 좌우회전과 배율변경 그리고 그 조합을 찾아낼 수 있고 그 값을 정량적으로 표현할 수 있다.[4] 그러나 계산시간이 많이 소요되고 일련의 프레임들이 한 개의 이동물체에 의해 가려질 경우에는 좌우회전이 잘못 검출된다는 단점이 있다.

Tomography[5]와 Visual Icon[6] 기술에서는 비디오 콘텐츠 파싱(contents parsing)을 위해서 보다 더 정교한 카메라 움직임 검출 알고리즘을 사용한다. 그러나 이 방법에서 구할 수 있는 매개변수는 좌우회전과 상하회전, 배율변경 등 몇가지로 제한되어 있다.

보다 최근의 연구에서는 특징점(feature point)과 대응점간의 외극기하적 제한조건(epipolar constraints)에 의해 유도되는 E-행렬[7] 혹은 F-행렬[8]의 해를 구하는 방법이 주류를 이루고 있다.

특징점으로는 전형적으로 모서리나 직선을 사용하며, 최근의 연구들[9,10,11]에서도 비교적 특징점을 찾기 쉬운 건물, 탱크 등의 인공물을 대상으로 실험을 수행하였다. 그

러나 실제의 영상에서는 그러한 인위적인 도형이 존재하지 않는 경우가 많기 때문에 에러에 민감하다. 이러한 에러 민감성을 해결하기 위해서 과평가(over determined)기법을 사용하지만, 추정한 대응점 쌍의 오차분포가 영평균오차(zero mean error)가 아닐 경우에는 추정에 실패한다. 그리고 영상에 따라 외극기하적 제한조건이 성립되지 않는 경우도 발생할 수 있다.

해를 구하는 방법은 선형해법과 비선형해법으로 분류할 수 있다. 선형 최소자승법의 경우에는 E-행렬의 구축에 사용되는 특징점과 대응점 쌍들이 독립적이지 않은 경우에는 퇴화차수(nullity)가 증가하여 적용이 곤란해지는 등의 문제점이 있다.

비선형 해법에서는 흔히 수치적 반복법을 사용하는데 계산량이 방대해진다는 단점이 있다. 그리고 정확한 초기값 가정을 하지 못할 경우에는 해가 부정확하거나 수렴하지 않는 경우도 발생한다.[12]

직접법은 순수한 이동, 순수한 회전, 혹은 회전을 아는 경우의 이동값 등 제한된 조건에서만 적용이 가능하다는 단점이 있다.

광류(optical flow)기술에서는 유동장(flow field)이 지역적으로 부드러워야 한다는 가정이 성립되어야 적용이 가능한데, 대부분의 실제 영상에서는 다른 움직임을 취하는 물체들이 여러 개 존재하기 때문에, 움직임 필드는 균질(homogeneous)하지 않다. 이러한 경우에는 대부분의 움직임 분석이 실패하곤 한다.

3장에서는 본 연구에서 제안하는 기술을 소개하고, 4장에서 실험결과를 통하여 제안된 기술의 유용성을 입증한다.

## III. 카메라 움직임 추정 기술

### 1. 원근투영모델에서의 기본관계식 유도

원근투영모델(perspective projection)에서 3차원 실세계 공간상의 점 P는 그림 1과 같이 2차원 영상면의 점 p에 투영된다.

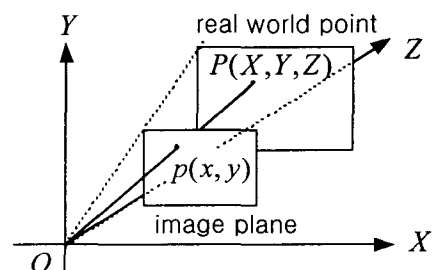


그림 1. 원근투영모델

원근 투영에서의 이 p와 P, 두 점은 그림 2와 같은 기하적 관계를 갖게 된다. 여기서 f는 초점거리이고 λ는 f와 Z의 비이다.

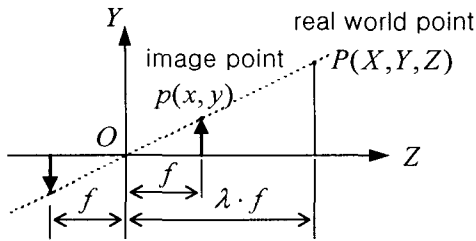


그림 2. 원근투영기하

영상 프레임의 중심에 x-y 평면의 원점이 위치한다고 가정 하자. 그리고 이 원점에서 그림 3과 같이 3차원 왼손 좌표계를 가상하고  $\theta$ ,  $\Psi$ 와  $\phi$ 를 각각 X, Y, Z축에 대한 카메라 회전각도라 하고,  $T_x$ ,  $T_y$ 와  $T_z$ 를 각각 X, Y, Z축에 대한 카메라 이동거리라고 하자. 그리고  $s$ 를 카메라의 배율(zooming factor)이라 하자.

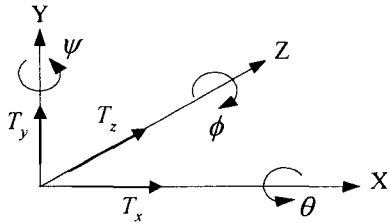


그림 3. 카메라 움직임 매개변수

만약 카메라 움직임이 발생한다면, X-Y평면에 대한 수직거리 즉 깊이가  $z$ 인 어느 투영점  $p(x, y)$ 는 그림 4와 같이 새로운 좌표  $p'(x', y')$ 로 이동할 것이다.

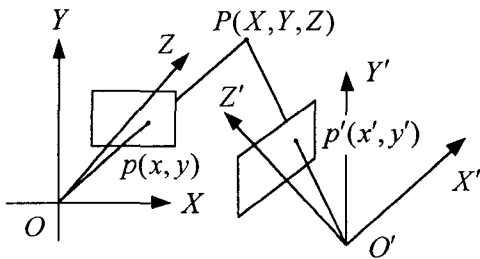


그림 4. 카메라 좌표 변환

이때  $x'$ 와  $y'$ 는 식(1)과 같이 유도할 수 있다.

$$\begin{aligned}
 x' &= f' \frac{(\lambda x - \phi \lambda y + \phi \lambda f + T_x)}{-\phi \lambda x + \theta \lambda y + \lambda f + T_z} \\
 &= \frac{f'}{f} \frac{x - \phi y + \phi f + f \frac{T_x}{Z}}{1 - \frac{\phi}{f} x + \frac{\theta}{f} y + \frac{T_z}{Z}} \\
 y' &= f' \frac{(\phi \lambda x + \lambda y - \theta \lambda f + T_y)}{-\phi \lambda x + \theta \lambda y + \lambda f + T_z}
 \end{aligned} \quad (1)$$

$$= \frac{f'}{f} \frac{\phi x + y - \theta f + f \frac{T_y}{Z}}{1 - \frac{\phi}{f} x + \frac{\theta}{f} y + \frac{T_z}{Z}}$$

식(1)에서, 카메라 회전과 배율변경에 의한 투영점의 변위는 깊이  $Z$ 에 무관하고, 카메라의 이동에 의한 투영점의 변위는 깊이  $Z$ 에 반비례한다는 사실을 알 수 있다.

식이 깊이  $Z$ 에 독립적이 되도록 하기 위해서, 우리는 깊이  $Z$ 가 무한대에 가까운 원경의 점을 배경으로 간주한다. 그리고 카메라 움직임에 독립적으로 움직이는 개개의 물체와 근경은 전경으로 간주한다.

만약  $\phi \ll 1$ ,  $\theta \ll 1$ 이고  $f \ll x$ ,  $f \ll y$ 라고 가정하면, 식(1)은 식(2)와 같이 간략화된다.

$$\begin{aligned}
 x' &\approx s(x - \phi y + \phi f) \\
 y' &\approx s(y + \phi x - \theta f)
 \end{aligned} \quad (2)$$

점  $p(x, y)$ 의 움직임 벡터의  $x, y$  방향 성분을 각각  $v_x$ 와  $v_y$ 라고 표기하면, 식(2)로부터 식(3)과 같이 움직임 벡터와 카메라 동작 매개변수간의 기본적인 관계식을 확립할 수 있다.

$$\begin{aligned}
 v_x &= x' - x = sx - x - s\phi y + s\phi f \\
 v_y &= y' - y = sy - y + s\phi x - s\theta f
 \end{aligned} \quad (3)$$

## 2. 움직임 벡터의 대칭관계

우리는 임의의 점  $p(x, y)$ 에 대하여, 원점에 대칭인 점, Y축에 대칭인 점,  $+90^\circ$  회전한 점,  $-90^\circ$  회전한 점등과 같이 기하학적으로 특수한 관계인 4점에 주목한다. 이러한 점들을 그림 5에서와 같이  $p(x_1, y_1)$ ,  $p(x_2, y_2)$ ,  $p(x_3, y_3)$ 와  $p(x_4, y_4)$ 로 표기하자.

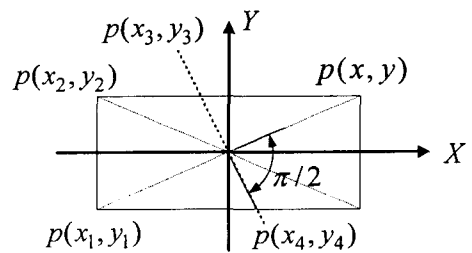


그림 5. 영상평면에서 점  $p(x, y)$ 와 기하적으로 특수관계인 4개의 점

만약 점  $p(x, y)$ 와 특수관계인 점들이 모두 원경에 속한다면, 카메라 회전의 종류에 따라 점  $p(x, y)$ 의 움직임 벡터와 특수관계인 점들의 움직임 벡터는 서로 크기가 같으면서 방향이 반대이거나, 크기와 방향이 모두 같은 등의 특수한 관계가 성립된다.

그림 5의 원점에 대칭인 두 점  $p(x, y)$ ,  $p(x_1, y_1)$ 에서 카메라 동작에 의해 야기되는 움직임 벡터의 X좌표축과 Y좌표축 방향 성분간에는 그림 6과 같이  $v_\phi = v_{1\phi}$ ,  $v_\theta = v_{1\theta}$ ,  $v_z = -v_{1z}$ 인 관계가 성립된다.

그림 6에서  $v_\psi$ 와  $v_{1\psi}$ 는 좌우회전에 의해 야기되는 벡터의 X축 방향 성분을 나타내고,  $v_\theta$ 와  $v_{1\theta}$ 는 상하회전에 의한 Y축 방향 성분,  $v_z$ 와  $v_{1z}$ 는 배율변경에 의한 Z축 방향 성분을 나타낸다.

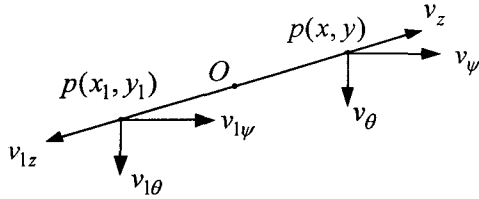


그림 6. 점  $p(x, y)$ 와 점  $p(x_1, y_1)$ 의 움직임 벡터간의 대칭관계

그림 5의 Y축에 대칭인 두 점  $p(x, y)$ ,  $p(x_2, y_2)$ 에서 카메라 동작에 의해 야기되는 움직임 벡터의 X좌표축과 Y좌표축 방향 성분간에는 그림 7과 같이  $v_\psi = v_{2\psi}$ ,  $v_\theta = v_{2\theta}$ 의 관계가 성립되고,  $v_z$ 와  $v_{1z}$ 는 Y축에 대칭이다.

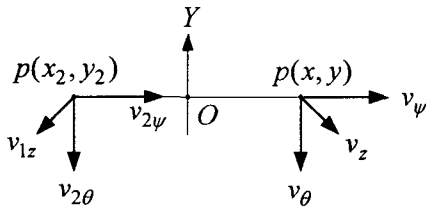


그림 7. 점  $p(x, y)$ 와 점  $p(x_2, y_2)$ 의 움직임 벡터간의 대칭관계

그리고 그림 5의 점  $p(x, y)$ 를  $\phi/2$ 도 회전한 점  $p(x_4, y_4)$ 에서 카메라 동작에 의해 야기되는 움직임 벡터의 각 좌표축 방향 성분간에는 그림 8과 같이  $v_\psi = v_{4\psi}$ ,  $v_\theta = v_{4\theta}$ 의 관계가 성립되고,  $v_z$ 와  $v_{4z}$ 는 서로 직각을 이룬다.

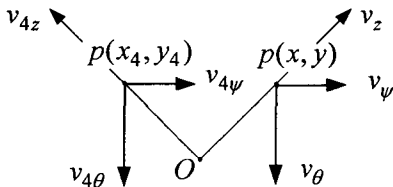


그림 8. 점  $p(x, y)$ 와 점  $p(x_4, y_4)$ 의 움직임 벡터간의 대칭관계

### 3. 배경 판별 및 카메라 매개변수 추정

영상 프레임에서 임의의 한 점과 이 점에 대하여 특수한 관계인 점들의 움직임 벡터들이 위 그림 6, 7, 8과 같은 조건을 만족하는 지 검토하면 그 점이 배경에 속하는지 아닌지를 판별할 수 있고, 배경에 속하는 점들의 변위를 조사하면 카메라 움직임 매개변수를 추정할 수 있다. 그런데 잡

음 혹은 동체의 움직임에 의해 우연히 두 점이 조건을 만족하는 경우도 있을 수 있기 때문에 보다 정교하게 3점간의 관계를 동시에 검증하는 판별식과 추정식을 유도한다.

또한 X-Y좌표평면의 4사분면의 대칭성을 이용하면, 영상의 모든 점에 대하여 특수관계인 점들을 아래의 4가지 트리플(triple) 유형 중 한가지에 속하도록 분류할 수 있다. 트리플 유형1은 그림 5에서의 세 점,  $p(x, y)$ ,  $p(x_1, y_1)$ ,  $p(x_2, y_2)$ 로 구성되고, 트리플 유형2는 세 점,  $p(x, y)$ ,  $p(x_1, y_1)$ ,  $p(x_3, y_3)$ 으로 구성된다. 트리플 유형3은  $p(x, y)$ ,  $p(x_2, y_2)$ ,  $p(x_3, y_3)$ 으로 구성되고, 트리플 유형4는 세 점,  $p(x, y)$ ,  $p(x_2, y_2)$ ,  $p(x_4, y_4)$ 로 구성된다.

각 트리플 유형에 대하여, 주어진 점이 원경에 속하는지를 판별하는 방정식 집합과 카메라 움직임 매개변수를 추정하는 방정식 집합을 유도한다.

#### 1) 트리플 유형 1

트리플 유형 1은 세 점,  $p(x, y)$ ,  $p(x_1, y_1)$ ,  $p(x_2, y_2)$ 으로 구성된다.

$p(x_1, y_1) = p(-x, -y)$  이기 때문에 식(3)에 의하면, 점  $p(x_1, y_1)$ 와 점  $p(x', y')$ 간의 움직임 벡터의  $x, y$  성분은 식(4)와 같이 표기할 수 있다.

$$v_{x1} = x'_1 - x_1 = -sx + x + s\phi y + s\psi f \tag{4}$$

$$v_{y1} = y'_1 - y_1 = -sy + y - s\phi x - s\psi f\theta$$

또한 위 두 점의 움직임 벡터의  $x, y$  방향 성분의 합과 차를 아래 식(5)와 같이 정의하여 구할 수 있다.

$$v_{x1}^+ = v_x + v_{x1} = 2s\psi f \tag{5}$$

$$v_{y1}^+ = v_y + v_{y1} = -2s\psi f\theta$$

$$v_{x1}^- = v_x - v_{x1} = 2(s-1)x - 2s\phi y$$

$$v_{y1}^- = v_y - v_{y1} = 2(s-1)y + 2s\phi x$$

같은 표기법을 사용하여 점  $p(x, y)$ 와 점  $p(x_2, y_2)$ 의 움직임 벡터의 합과 차에 관한 식을 아래 식(6)과 같이 유도할 수 있다.

$$v_{x2}^+ = v_x + v_{x2} = 2s(-\phi y + f\psi) \tag{6}$$

$$v_{y2}^+ = v_y + v_{y2} = 2(sy - y - s\psi f\theta)$$

$$v_{x2}^- = v_x - v_{x2} = 2(s-1)x$$

$$v_{y2}^- = v_y - v_{y2} = 2s\phi x$$

그리고 점  $p(x_1, y_1)$ 과 점  $p(x_2, y_2)$ 에 관한 식도 역시 같은 방법으로 식(7)과 같이 유도할 수 있다.

$$v_{x12}^+ = v_{x1} + v_{x2} = 2(-sx + x + s\psi f) \tag{7}$$

$$v_{y12}^+ = v_{y1} + v_{y2} = -2s(\phi x + f\theta)$$

$$v_{x12}^- = v_{x1} - v_{x2} = 2s\phi y$$

$$v_{y12}^- = v_{y1} - v_{y2} = -2(s-1)y$$

식(6)과 (7)에서 벡터들은 불럭정합에 의해 구할 수 있는 기저수이므로 정리하면, 미지수인 배율  $s$ 를 식(8)과 같이 구할 수 있다.

$$s = 1 + \frac{v_{x2}^-}{2x} = 1 - \frac{v_{y12}^-}{2y} \quad (8)$$

판별식 집합은 아래 식(9)와 같이 유도된다. 이 판별식들은 세 점 중에서 어떠한 한 점이 식(9)를 만족하지 않으면, 최소한 그들 세 점 중 한 점이 배경에 속하지 않는다는 것을 의미한다.

$$xv_{y12}^- = -yv_{x12}^- \quad (9)$$

$$v_{x1}^+ = v_{x2}^- + v_{x12}^+$$

$$v_{y1}^+ = v_{y2}^- + v_{y12}^+$$

$$yv_{y2}^- = xv_{x12}^-$$

식(4)에서 (9)를 사용하여, 아래와 같이 카메라 움직임 매개변수를 추정하는 식(10)을 유도할 수 있다.

$$f\phi = \frac{xv_{x1}^+}{2x + v_{x2}^-} \quad (10)$$

$$f\theta = \frac{-yv_{y1}^+}{2y - v_{y12}^-}$$

$$\begin{aligned} \phi &= \frac{v_{y2}^-}{2sx} = \frac{v_{x12}^-}{2sy} \\ &= \frac{v_{y2}^-}{2x + v_{x2}^-} = \frac{v_{x12}^-}{2y - v_{y12}^-} \end{aligned}$$

식의 계산에서 점(x,y) 및 점(x',y')의 위치, 초점거리 f, 벡터 (v<sub>x</sub>,v<sub>y</sub>)의 값들은 pel로써 나타낸다. 2차원 영상에서 3차원 실세계 정보를 복원하는 것은 "ill-posed problem"에 속하며, 스케일(scale)에 관한 사진 정보가 주어지지 않은 임의의 영상의 경우 비례계수의 범위 내(up to scale factor)에서 구할 수밖에 없음은 잘 알려져 있다.[12]

본 연구에서는 각 회전각을 초점거리 f 혹은 영상의 폭 W와의 관계식 형태로 구한다. 즉, 위 식에서 f 와 W는 비례계수의 역할을 한다. 좌우회전각  $\Psi$ , 상하회전각  $\Theta$ 는 실제 영상에 나타나는 효과를 나타내는  $\Psi=f\phi$  와  $\Theta=f\theta$  의 형태로 구한다. 정면회전각  $\Phi$ 도 영상에 나타나는 효과를 나타내는  $\Phi=\phi W/2$ 의 형태로 구한다. 영상의 폭을 W라 할 때,  $\Phi$ 는 정면회전이 영상좌표값 (W/2, 0)에 미치는 효과의 pel 값을 나타낸다. 배율변경의 효과도 영상에 나타나는 효과를 나타내는  $S=(s-1)W/2$ 의 형태로 계산한다.

## 2) 트리플 유형 2

트리플 유형 2는 세 점,  $p(x, y), p(x_1, y_1), p(x_3, y_3)$ 으로 구성된다. 위 유형 1에서와 같은 표기법을 사용하여, 트리플 유형 2에 관한 카메라 매개변수 추정식을 식(11)과 같이 유도할 수 있다.

$$s = 1 + \frac{xv_{x1}^- + yv_{y1}^-}{2(x^2 + y^2)} \quad (11)$$

$$\phi = \frac{xv_{y1}^- - yv_{x1}^-}{2(x^2 - y^2) + xv_{x1}^- - yv_{y1}^-}$$

$$f\phi = \frac{(x^2 + y^2)v_{x1}^+}{2(x^2 + y^2) + xv_{x1}^- + yv_{y1}^-}$$

$$f\theta = -\frac{(x^2 + y^2)v_{y1}^+}{2(x^2 + y^2) + xv_{x1}^- + yv_{y1}^-}$$

그리고 배경판별식은 식(12)와 같이 유도할 수 있다.

$$v_{x3}^- = -v_{y13}^- \quad (12)$$

$$v_{y3}^+ = v_{x13}^+$$

$$v_{y3}^- = v_{x13}^-$$

$$v_{x3}^- = -v_{y13}^-$$

## 3) 트리플 유형 3

트리플 유형 3은 세 점,  $p(x, y), p(x_2, y_2), p(x_3, y_3)$ 으로 구성된다. 위 유형 1에서와 같은 표기법을 사용하여, 트리플 유형 3에 관한 카메라 매개변수 추정식을 식(13)과 같이 유도할 수 있다.

$$s = 1 + \frac{v_{x2}^-}{2x} \quad (13)$$

$$\phi = \frac{v_{y2}^-}{2x + v_{x2}^-}$$

$$f\phi = \frac{xv_{x2}^+ + yv_{y2}^-}{2x + v_{x2}^-}$$

$$f\theta = \frac{yv_{x2}^- - xv_{y2}^+}{2x + v_{x2}^-}$$

그리고 배경판별식은 식(14)와 같이 유도할 수 있다.

$$(x^2 + y^2)v_{x2}^- = x(x + y)v_{x3}^- - x(x - y)v_{y3}^- \quad (14)$$

$$xv_{x2}^+ + yv_{y2}^- = x(v_{x3}^+ + v_{y3}^+)$$

$$(x - y)v_{x2}^- + x(v_{x23}^- + v_{y23}^-) = 0$$

$$(x + y)v_{y2}^- + x(v_{x23}^+ + v_{y23}^+) = 0$$

## 4) 트리플 유형 4

트리플 유형 4는 세 점,  $p(x, y), p(x_2, y_2), p(x_4, y_4)$ 로 구성된다. 카메라 매개변수 추정식은 식(15)와 같이 유도할 수 있다.

$$s = 1 + \frac{v_{x2}^-}{2x} \quad (15)$$

$$\phi = \frac{v_{y2}^-}{2x + v_{x2}^-}$$

$$f\psi = \frac{xv_{x2}^+ + yv_{y2}^-}{2x + v_{x2}^-}$$

$$f\theta = \frac{yv_{x2}^- - xv_{y2}^+}{2x + v_{x2}^-}$$

그리고 판별식은 아래 식(16)과 같이 유도된다.

$$\begin{aligned} (x^2 + y^2)v_{x2}^- &= x(x - y)v_{x4}^- + x(x + y)v_{y4}^- \\ yv_{x2}^- - xv_{y2}^+ + x(v_{x4}^+ + v_{y4}^-) &= 0 \\ (x + y)v_{x2}^- + x(v_{x4}^- - v_{y4}^-) &= 0 \\ (x + y)v_{y2}^- + x(v_{x4}^- + v_{y4}^-) &= 0 \end{aligned} \quad (16)$$

#### IV. 실험결과

이 장에서는, 전역움직임을 보상한 프레임에 대한 직관적인 관찰, 그리고 추정된 카메라움직임 매개변수들에 대한 수치적인 검토를 통해서, 제안된 기술을 입증한다.

제안된 기술의 증명을 위한 실험은 다음과 같이 6단계로 수행된다. 첫째, 연속하는 두 프레임간의 움직임벡터를 추정한다. 둘째, 제안된 판별식에 의해 이 벡터가 배경에 속하는지 아닌지를 판별한다. 만약 속한다면, 이 벡터들은 카메라움직임 추정 방정식들에게 보내진다. 셋째, 카메라움직임 추정 방정식들이 카메라움직임 매개변수 set을 생성한다. 넷째, 추정된 각각의 매개변수 값의 정확성을 검증하기 위하여 지정된 범위 내에서 반 픽셀 단위로 변경한 매개변수 set ( $\theta \pm 2.5$ ,  $\psi \pm 2.5$ ,  $\phi \pm 2.5$ ,  $z \pm 2.5$ )들을 적용한 14,641개의 보상 프레임들을 만든다. 다섯째, 이 보상된 프레임들을 참조프레임과 비교한 차분값이 가장 작은 것을 찾아낸다. 보상 프레임과 참조 프레임을 비교할 때 움직이는 물체가 있는 영역은 제외시켜 움직이는 물체에 의해 차분값이 발생하는 것을 억제한다. 여기서 가장 작은 밝기차분값(intensity difference)을 생성한 매개변수 set은 가장 실제에 가까운 매개변수 set을 의미한다. 여섯째, 가장 작은 차분값을 생성한 매개변수 set과 제안된 기술이 추정한 매개변수 set을 비교한다. 모든 영상열들의 모든 프레임이 처리될 때까지 이러한 과정을 반복한다.

실험에서는 영상처리분야에서 "잘 알려진 영상열(well known sequence)"로 불리는 'Football', 'Stefan', 'Mobile', 'Flower Garden', 'Coast', 'Children' 등을 사용하였다.

그림 9(a)에서부터 9(f)는 'Football' 영상열에 대하여 수행한 실험의 결과를 보여준다. 그림 9(a)와 9(b)는 프레

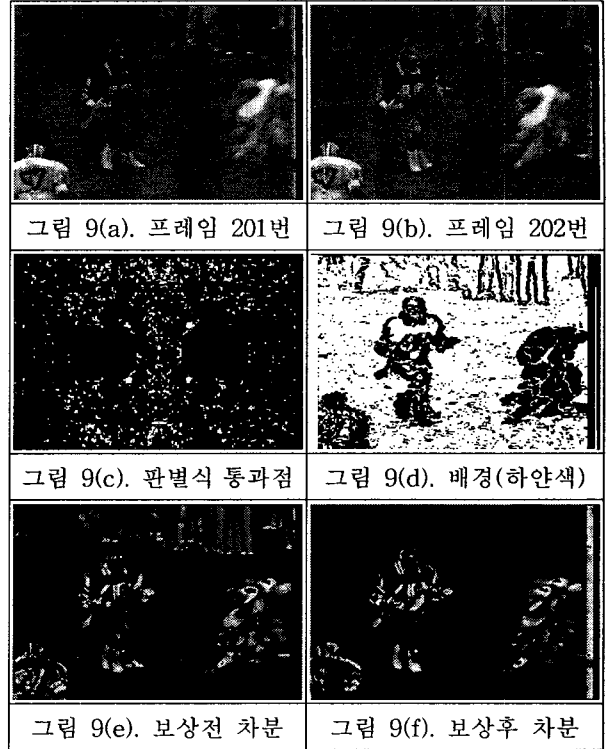


그림 9. Football 영상열 실험결과

임 #201과 프레임 #202 이다. 그림 9(c)는 이진 영상인데, 하얀 부분은 배경판별식을 통과한 모든 트리플 점들을 나타낸다. 그림에서 움직이는 물체에 해당하는 점들은 거의 모두 전경에 속하기 때문에 판별식을 통과하지 못하는 것을 확인할 수 있다. 그림 9(d)는 프레임 #201의 전역움직임 보상한 영상과 프레임 #202와의 차분이 문턱값(threshold) 5 이하인 점들을 나타낸 영상인데 배경판별식에 의해 분리된 배경과 전경을 보여주고 있다. 여기서 하얀 부분은 배경을 나타내고, 검은 부분은 전경을 나타낸다.

그림 9(e)는 연속하는 두 프레임간의 밝기차이를 나타낸 것이다. 그림 9(f)는 같은 프레임의 전역움직임 보상 프레임이다. 이 그림은 카메라에 의해 도입된 움직임은 보상되고 물체의 독립적인 움직임은 그렇지 않음을 보여준다. 겹쳐지지 않는 영상경계는 처리하지 않았다. 두 프레임 사이의 짧은 시간 간격 동안에 실제 거의 움직이지 않았을 배경(정지한 관중포함)의 카메라동작유도움직임은 보상되어 사라지고 카메라에 관계없이 자의적으로 크게 움직이는 선수들의 움직임에 의한 차이값만 결과의 영상에 나타나 있다.

그림 10(a)에서부터 10(f)는 'Stefan' 영상열에 대하여 수행한 실험의 결과를 보여준다. 그림 10(a)와 10(b)는 프레임 #8과 프레임 #9 이다. 그림 10(c)는 배경판별식을 통과한 점들을 나타낸다. 그림 10(d)는 배경판별식에 의해 분리된 배경과 전경을 보여주고 있다. 하얀 부분은 배경을 나타내고, 검은 부분은 전경을 나타낸다. 그림 10(e)는 연속하는 두 프레임간의 밝기 차이를 나타낸 것이다. 그림 10(f)는



그림 10. Stefan 영상열 실험결과

같은 프레임의 전역움직임 보상 프레임이다. 이 그림은 카메라에 의해 도입된 배경의 움직임은 보상되고 동체(테니스 선수)의 독립적인 움직임은 보상되지 않음을 보여준다.

각기 10장의 프레임으로 구성된 8개의 영상열에 대한 실험을 수행하여 총 320쌍의 매개변수를 산출하였다. 8가지의 영상열 중에서 다양한 카메라 회전 이 내재된 5가지 영상열에 대한 실험결과를 표 1, 2, 3, 4, 5에 요약한다. 표의 각 칸에서, 왼쪽 수치는 본 논문에서 제안한 기술에 의해 추정된 값이고 오른쪽 수치는 전역검색한 값이다. 표에 제시한 바와 같이, 거의 대부분의 추정값들이 전역검색한 값들과 완전히 일치하였고, 0.5 화소 거리 보다 더 큰 차이가 나는 것은 없었다. 이는 실험결과와 정확성을 입증한다.

이와 같은 수치적인 비교 외에, 직관에 의해서도 실험 결과의 정확성을 쉽게 확인할 수 있다. 표 1과 표 2에 제시된 Football과 Stefan 영상열의 좌우회전각 추정값이 다른 영상열의 추정값에 비해 매우 크다. 실제로 두 영상열을 시각에 의해 관찰해보면 빠르게 이동하는 운동선수를 추적하기 위해 카메라가 다른 영상열에 비해 좌우로 크게 회전했다는 사실을 확인할 수 있다. 또한 표 5의 Children 영상열의 경우 추정된 값이 전부 0이다. 이 영상열을 화면에서 작동시켜보면 카메라 움직임이 전혀 없다는 사실을 시각적으로 쉽게 확인할 수 있다.

이상의 실험결과에서 분석한 바와 같이, 제안한 기술은

표 1. Football 영상열의 카메라 회전각

	Tilting	Panning	Rolling	Zooming
#1~#2	-0.5:-0.5	-8.5:-9.0	0.0:0.0	0.0:0.0
#2~#3	0.0:0.0	-8.5:-9.0	0.0:0.0	0.0:0.0
#3~#4	0.0:0.0	-8.5:-9.0	0.0:0.0	0.0:0.0
#4~#5	-0.5:-0.5	-8.0:-8.0	0.0:0.0	0.0:0.0

표 2. Stefan 영상열의 카메라 회전각

	Tilting	Panning	Rolling	Zooming
#1~#2	0.5:0.5	-7.5:-7.5	0.0:0.0	-2.0:-2.0
#2~#3	0.5:0.5	-7.5:-7.5	0.0:0.0	-2.0:-2.0
#3~#4	0.5:0.5	-7.0:-7.0	-0.5:-0.5	-2.5:-2.5
#4~#5	0.5:0.5	-6.5:-6.5	0.0:0.0	-2.5:-2.5

표 3. Mobile 영상열의 카메라 회전각

	Tilting	Panning	Rolling	Zooming
#1~#2	-0.5:-0.5	0.5:0.5	0.5:0.5	-0.5:-0.5
#2~#3	-0.5:-0.5	0.5:0.5	0.5:0.5	-0.5:-0.5
#3~#4	-0.5:-0.5	0.5:0.5	0.5:0.5	-0.5:-0.5
#4~#5	-0.5:-0.5	0.5:0.5	0.5:0.5	-0.5:-0.5

표 4. Coast 영상열의 카메라 회전각

	Tilting	Panning	Rolling	Zooming
#1~#2	0.0:0.0	1.5:1.5	0.0:0.0	0.0:0.0
#2~#3	0.0:0.0	1.5:1.5	0.0:0.0	0.0:0.0
#3~#4	0.0:0.0	1.5:1.5	0.0:0.0	0.0:0.0
#4~#5	0.0:0.0	1.5:1.5	0.0:0.0	0.0:0.0

표 5. Children 영상열의 카메라 회전각

	Tilting	Panning	Rolling	Zooming
#1~#2	0.0:0.0	0.0:0.0	0.0:0.0	0.0:0.0
#2~#3	0.0:0.0	0.0:0.0	0.0:0.0	0.0:0.0
#3~#4	0.0:0.0	0.0:0.0	0.0:0.0	0.0:0.0
#4~#5	0.0:0.0	0.0:0.0	0.0:0.0	0.0:0.0

주) 매개변수표기: 추정값 대 전역검색값

카메라의 움직임과는 관계없이 자의적으로 움직이는 동체에 해당하는 화소를 판별하여 계산에서 배제할 수 있기 때문에, 여러 개의 동체가 빠르게 움직이는 Football과 같은 영상에서도 결과가 정확하다. 기존 E-행렬 방식 기술의 경우, 제한조건하의 최소자승적 풀이를 위해 SVD(Singular Value Decomposition)를 사용하는 경우의 성능이  $O(n^3)$  입에 비하여, 제안한 기술은 선형방정식에 움직임벡터를 한번 적용하기 때문에  $O(n)$ 의 성능을 갖는다. 또한 인공적인 모서리와 같은 특징점이 없는 "잘 알려진" 영상열에서 수행된 실험결과로 볼 때, 제안하는 기술의 광범위한 적용 범위와 응용가능성을 알 수 있다.

## V. 결론

본 논문에서는 외극기하적 제한조건에 의존하지 않고 영상 프레임에서 기하학적으로 대칭인 세 점의 움직임 벡터의 특수한 관계를 이용하여 카메라의 움직임을 추정하는 기술을 제안하였다.

제안된 기술은 물체의 깊이 혹은 카메라 움직임의 종류 등과 같은 사전 정보를 필요로 하지 않고 오직 연속된 두 프레임 그 자체만의 정보를 사용한다. 그리고 제안된 선형 방정식을 움직임 벡터에 단 한번 적용하여, 좌우회전, 상하회전, 정면회전 그리고 배울변경과 같은 여러 가지 카메라 움직임 매개변수들을 한꺼번에 산출하며, 동체를 구별하는 판별식을 사용하여 배경정보만으로 카메라 움직임추정을 한다. 또한 기존 연구들과는 달리 수치해석이나 반복법을 사용하지 않기 때문에 계산능력은  $O(n)$ 으로서 화면 크기에 선형 비례한다. 따라서 계산이 보다 정확하고 실시간처리가 가능하다. 다양한 종류의 영상에 대한 실험결과를 제안된 기술의 타당성과 광범위한 응용가능성을 보여준다.

접수일자 : 2002. 3. 10      수정완료 : 2002. 4. 19

본 연구는 한국과학재단 목격기초연구 (과제번호: 104533) 지원으로 수행되었음.

### 참고문헌

[1] ISO/IEC JTC1/SC29/WG11 N2459, "Overview of the MPEG-4 Standard," Atlantic City, Oct. 1998.

[2] ISO/IEC JTC1/SC29/WG11 2702, "Description of automatic segmentation techniques developed and tested for MPEG-4 version1," Oct. 1997.

[3] H. J. Zhang, A. Kankanhalli and S. W. Smoliar, Automatic partitioning of full-motion video, Multimedia Systems, ACM-Springer 1, 1, 1993, 10-28

[4] Y. T. Tse and R. L. Baker, "Global zoom/pan estimation and compensation for video compression," in Proc. ICASSP, pp. 2725-2728, 1991.

[5] A. Akutsu and Y. Tonomura, Video tomography: An efficient method for camerawork extraction and motion analysis, Proc. ACM Multimedia Conference, San Francisco, Oct. 1993

[6] Y. Tonomura, A. Akutsu, K. Otsuji and T. Sadakata, VideoMAP and VideoSpaceIcon: Tools for anatomizing video content, Proc. InterChi'93, ACM (1994) pp.131-136, 1993

[7] T. S. Huang, "Determining three-dimensional motion and structure from two perspective views," Ch.14 in Handbook of Patt. Recog. and Image proc., Academic press, 1986

[8] O. Faugeras, T. Luong and S. Maybank, Camera self-calibration: theory and experiments, Proceedings of the 2nd European Conference on Computer Vision, pages 321-334, Santa Margherita, Italy, Springer-Verlag 1992

[9] S. Seitz, "The Space of All Stereo Images," Proc. IEEE 8th Intl Conf. Computer Vision, ICCV 2001, Vol. 1, pp. 26-33, Vancouver, Canada, July 2001

[10] R. Vidal, Y. Ma, S. Hsu, and S. Sastry, "Optimal Motion Estimation from Multiview Normalized Epipolar Constraint," Proc. IEEE 8th Intl Conf. Computer Vision, ICCV 2001, Vol. 1, pp. 26-33, Vancouver, Canada, July 2001

[11] C. Rother and S. Carlsson, "Linear Multi View Reconstruction and Camera Recovery," Proc. IEEE 8th Intl Conf. Computer Vision, ICCV 2001, Vol. 1, pp. 26-33, Vancouver, Canada, July 2001

[12] A. Murat Tekalp, Digital Video Processing, Prentice Hall, pp.158-165, 1995

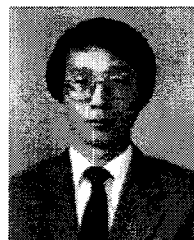
### 문성현 (Sung-Hun Moon)

正會員

1980년 영남대학교 기계공학과 (공학사)

1990년 영남대학교 전자공학과 (공학석사)

1993년 영남대학교 전자공학과 (박사수료)



1983년~1987년 (주)현대건설

1991년~현재 대구과학대학 컴퓨터공학과 부교수

관심분야 : 컴퓨터비전, 로보틱스, 네트워크

### 박영민 (Young-Min Park)

正會員

1993년 영남대학교 전산공학과 (공학사)

1995년 영남대학교 전산공학과 (공학석사)

2001년 영남대학교 컴퓨터공학과 (박사수료)



1995년~1999년 삼성전자(주) 정보통신연구소

2000년~현재 경운대학교 소프트웨어공학과 전임강사

관심분야 : 영상처리, ASIC설계, 컴퓨터구조

### 윤영우 (Young-Woo Yoon)

正會員

1972년 영남대학교 전자공학과 (공학사)

1974년 영남대학교 전자공학과 (공학석사)

1984년 영남대학교 전자공학과 (공학박사)



1988년~현재 영남대학교 컴퓨터공학과 교수

관심분야 : 컴퓨터비전, 멀티미디어시스템, 컴퓨터구조